Michael A. May
Laboratory Assistant - summer student
8/26/96

Working with HTML and PDF Documents

After working on the World Wide Web at Fermi Lab for three summers Mike Shea thought that it would be good to have documentation on what I have done, so I wrote this document to explain some of the "tricks" that I have picked up while working here. This document has three main sections: working with html (Hyper-Text Markup Language), pdf file conversion, and a troubleshooting section. The html section covers converting documents, gives an overview of the Linac Controls web sight, some advanced html code, and shell scripts. The pdf section gives a quick run though on how convert documents to pdf format and how to touch up these documents.

HTML

Converting documents to html is fairly straight forward. A program called "rtftohtml" which runs on the Macintosh does a good job of converting a document to the html format. The first step in the process of converting a document written on a program such as Microsoft Word or Claris Works is to save it as a rtf (Interchange Format) document. Once the rtf document has been made just drop it on the icon of the rtftohtml program. This program will insert the html tags and convert the tables and pictures within the document to pict formated files. These pict formated files must be converted

to the gif format so that these files can be seen on a web browser such as Netscape. A Macintosh program called "GraphicConverter" is one of the best software packages for converting graphics and quickly converts even numerous graphic files to just about anything including the gif format. Now the only thing that needs to be done is to put the document and all the pictures with it on the server (this is covered in the updating section).

Overview

The documentation for the World Wide Web is organized under three main categories: Linac, Booster, and IRM. All three of these directories can be found in the www directory. Linac was the first directory created and is the largest. It cantains the image conf (also call image map) documents used in making images that have multiple links for all three of the main directories. All three of these directories have whats call a symbolic link which is a link that calls up the home page (which is call welcome.html) for a directory by just typing in the basic path. For instance the URL for the Linac Home Page is http://www-linac.fnal.gov/linac_controls/. This greatly reduces the length of the path since the URL would be http://www-linac.fnal.gov/home/florian/www/linac/welcome.html. The symbolic link for Booster is /booster_controls/. The symbolic link for IRM is just /irm/. For a diagram of the layout of the Linac Controls directory structure see appendix A.

Updating

Updating documents on the web or adding new ones is fairly easy since the directory structure is already in place. If the document was made or converted on a Macintosh "fetch" is probably the best

program for copying files over to the server. When html documents are being transferred using fetch, transfer the data as text. Transfer image files as raw data. If the document that needs to go on the web is replacing a document that is already on the web then just make sure that the name of the new document is the same as the name as the older document, and use fetch to copy it over to the same directory in which the older document is located. In other words overwrite the older document. This saves times because it eliminates the possibility of having to change the links inside a document that was linked to the older document. If the document is not an update of an older document that is already on the web then some additional work must be done. First use fetch to copy the file into a directory that pertains to that file on the server. Now some html programming needs to be done. The book *Teach Yourself WebPublishing with HTML 3.0 in a Week* or a book like it should be picked up to learn how to program in html.

List Tags

I will not cover how to program in html in this document, but I will go over some of the more advanced html "tags" that are not well documented. One thing that I picked up that is not documented anywhere that I have seen is a tag that makes a list without putting a number or a black ball at the beginning of each item in the list. This tag is useful if one desires to have a little picture of a colored ball at the beginning of each line instead of the black ball as seen throughout the Linac controls web sight. This tag that starts the list is <DL> and then a <DT> tag is put in the front of each item. The list is then ended with a </DL> tag.

Frames

It is also difficult to find documentation on frames. Frames are used on the Linac and Booster Home page. Frames are several documents that are all stuck on one page. Frames can be a little tricky to use, but hopefully this text will clear things up. First a main page is set up that calls all the other pages (frames). The pages are called up with the <FRAME> tag, but first the <FRAMESET> tag needs to be given to tell the html reader (usually a browser such as Netscape) that frames are going to be used. Within the <FRAMESET> tag the size for the area that each individual frame is going to use can be set. This is done by putting a "COLS" tag in for the percentage of a column that each frame is going to take or a "ROWS" tag for rows. After the <FRAMESET> tag the <FRAME> tags are put in. These tags need to have the path for the html document that they will bring up in that frame. It is also good to give each frame a name by putting 'NAME="(name of frame)"' inside the <FRAME> tag. After all the <FRAME> tags have been entered then a </FRAMESET> tag need to be entered. The example below is from the Linac controls home page.
<HTML>
<TITLE>Fermilab Linac Controls</TITLE>

<FRAMESET COLS="60%,40%">

    <FRAME NAME="Frame1" SRC="http://www-linac.fnal.gov/cgi-bin/LinacFrame">

    <FRAME NAME="Frame2" SRC="http://wwwlinac.fnal.gov/linac_controls/LinacFrame2.html">

</FRAMESET>

</HTML>

The COLS="60%,40%" give frame1 60% of the browser window and frame2 40%.

The next step in making a page that uses frames is to make the individual frames themselves. The only difference between a html document that is used as a frame and a regular html document are the links. A link can be set up so that it opens whatever it is linked to in the frame that it is in or any of the other frames on the page or it can open the document it is linked to on an entirely new page. A link that opens a document into the frame that it is located is the same as a regular link. Don't change anything. For the other links, however, there are some things that need to be done. When the line target="(name of frame)" is added to the tag of a link it brings up the document in the frame that is named. When bringing up a document on its on frame and therefor getting rid of the frames target=_top is added to the tad of the desired link. The example below is of the second frame used in the Linac Controls Home Page.

```
<HTML>
<TITLE>Linac Frame 2</TITLE>

<BODY Background="/linac_controls/backgroundPatterns/tile.gif">
<CENTER><H3> Linac Controls Menu </H3></CENTER>
<DL>
<DT><IMG SRC="//adwww.fnal.gov/www/green_ball.gif"><A
HREF="http://www-linac.fnal.gov/cgi-bin/Dir2HTML"
target="Frame1">Overview</A>
<DT>List of information. -(Directory Structure)
<DT><IMG SRC="//adwww.fnal.gov/www/green_ball.gif"><A
HREF="http://wwwlinac.fnal.gov/linac_controls/hardware/
hardware.html" target="Frame1">Hardware</A>
<DT><IMG SRC="//adwww.fnal.gov/www/green_ball.gif"><A
HREF="http://wwwlinac.fnal.gov/linac_controls/software/
software.html" target="Frame1">Software</A>
<DT><IMG SRC="//adwww.fnal.gov/www/green_ball.gif"><A
HREF="http://www-linac.fnal.gov/irm/" target=_top>IRM</A>
```

```
<DT>Internet Rack Monitor
<DT><IMG SRC="//adwww.fnal.gov/www/green_ball.gif"><A
HREF="http://www-linac.fnal.gov/booster_controls/welcome.html"
target=_top>Booster</A>
<DT>HLRF Control System
</DL>
<HR>
<CENTER> <A HREF="//adwww.fnal.gov/" target=_top><IMG
SRC="/linac_controls/gifs/ADHomeSml.gif"></A></CENTER>
<HR>
<CENTER><A HREF="//www.fnal.gov/" target=_top>Fermi Lab</A>
Home Page</TH></CENTER>
<HR>
<PRE>
Comments
<A HREF="http://www-linac.fnal.gov/~florian/"
target=_top>florian</A>@pea.fnal.gov
</PRE>
<HR>
<PRE>
 **************************
 *   (C)  Copyright 1994    *
 *  Universities Research   *
 *    Association, Inc.      *
 *   All rights reserved.   *
 **************************
</PRE>

</HTML>
```

Shell Script

The first frame in the Linac Controls Home Page is a shell script (cgi-script). This is a program that is made on the Sun using Unix commands. The shell script for the first frame creates an html document with a ticker at the bottom that counts the number of people that have logged onto the web. The first part of the program gets the information from the log of the number of people that have logged onto the Linac Controls Home Page since it has had the /linac_controls/ path. Then the number of people that logged on

before the home page had that path are added to the number that was retrieved from log. Next math is done to brake down the total number into individual digits. Then the proper images are called up to display the ticker. The example below is a shell script called LinacFrame used in making the first frame in the Linac Controls Home Page.

```
#!/bin/sh

echo "content-type: text/html"   -- "sets text to html"
echo

cat << EOF
<HTML>
<BODY Background="/linac_controls/backgroundPatterns/tile.gif">
<TITLE>Linac Frame 1</TITLE>
<CENTER><APPLET code="Lester.class" codebase=/java_classes/
width=290 height=360>
<PARAM name="image" value="/linac_controls/Title.gif">
<blockquote>
<hr>                     -- "Bob Florian's Java Applet"
<em>
Your browser can't run 1.0 Java applets,
so here's a picture the program brings up:</em>
<p>
<img src=/linac_controls/Title.gif width=290 height=360>
<hr>
</blockquote>
</APPLET>
</CENTER>
    A view of the Fermilab Linac looking from
the High Energy (400 Mev) end. Click <A HREF="racks.html"
target=_top>here</A> to view
one System or <A
HREF="/linac_controls/hardware/gendescript/general.html"
target=_top>here</A> for a general description of the system.
<P>
<P>
<PRE>
EOF
```

```
COUNT=`cat /home/http/httpd-log.*|grep linac_controls|grep -v
bflorian|grep -v pea|grep -v garlic|awk '{print $1}'|uniq|wc|awk
'{print $1}'`            -- "Calls up and looks at the web log"
COUNTTOTAL=`expr $COUNT + 2489`
COUNTS1=`expr $COUNTTOTAL / 100000`
COUNTER1=`expr $COUNTTOTAL - $COUNTS1 \* 100000`
COUNTS2=`expr $COUNTER1 / 10000`
COUNTER2=`expr $COUNTER1 - $COUNTS2 \* 10000`
COUNTS3=`expr $COUNTER2 / 1000`
COUNTER3=`expr $COUNTER2 - $COUNTS3 \* 1000`
COUNTS4=`expr $COUNTER3 / 100`
COUNTER4=`expr $COUNTER3 - $COUNTS4 \* 100`
COUNTS5=`expr $COUNTER4 / 10`
COUNTER5=`expr $COUNTER4 - $COUNTS5 \* 10`
COUNTS6=$COUNTER5
```
-- **"Gets the individual digits in the number"**
```
echo "<H3>You are visitor number <IMG
SRC=\"/linac_controls/CounterGifs/${COUNTS1}.gif\"
ALIGN=MIDDLE><IMG
SRC=\"/linac_controls/CounterGifs/${COUNTS2}.gif\"
ALIGN=MIDDLE><IMG
SRC=\"/linac_controls/CounterGifs/${COUNTS3}.gif\"
ALIGN=MIDDLE><IMG
SRC=\"/linac_controls/CounterGifs/${COUNTS4}.gif\"
ALIGN=MIDDLE><IMG
SRC=\"/linac_controls/CounterGifs/${COUNTS5}.gif\"
ALIGN=MIDDLE><IMG
SRC=\"/linac_controls/CounterGifs/${COUNTS6}.gif\"
ALIGN=MIDDLE></H3>"
```
-- **"Calls the up the image that corresponds to the digit"**
```
cat << EOF
</PRE>
</BODY></HTML>
```
-- **"Ends the html document that the shell script creates"**

PDF Conversion

Making pdf documents is done in the same way that printing is
done. Pick the pdf print tool in the chooser and then print the

document or image that needs to be converted. Any program that can print is capable of being used to convert documents to pdf.

Touching Up PDF Documents

Once a pdf file is created there are a few thing that can be done to make that file more useful. Bookmarks and links are two of the most common methods used on Linac Documents. Both of these are made using the program Acrobat Exchange. Bookmarks are made by going to the page that the bookmark will be a link to and then going to the bookmark menu under edit or just typing "command B." Then type the name of the bookmark in the bookmark display to the left of the document. The bookmark display can be viewed by clicking on the second button from the left on the menu bar. Links are made by using the link tool that is under the tools menu. This tool can make links to other pdf documents or can open Netscape and go to a particular document on the web.

Putting PDF Documents on the Web.

Putting pdf documents on the web is similar to putting html documents on the web. Fetch can still be used to transfer pdf documents to the server. Make sure, however, that all pdf documents are transferred as raw data instead of text. Make sure that the catalogs for the Local Controls Software are updated as well when a Local Controls Software Document is updated.
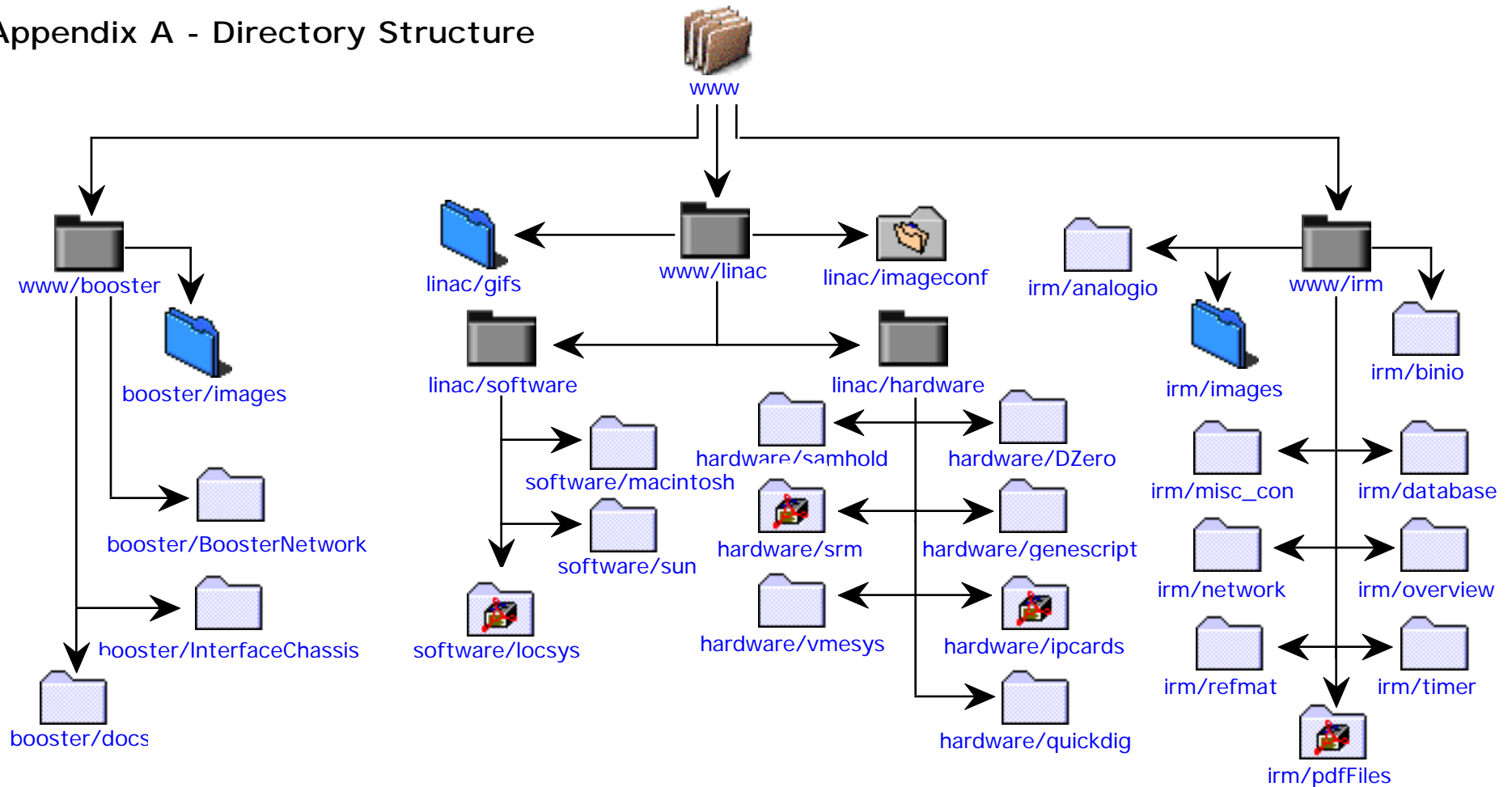
Troubleshooting

There are two applications that may cause some trouble that have been mentioned in this paper: the GraphicConverter and the pdf printer driver. The GraphicConverter is not System 7.5 friendly so it has a few errors on the Macintoshes used at Fermi since almost all of

them use System 7.5. The only thing that can be done is to get the upgrade as soon as it comes out. The current version of the GraphicConverter is 2.4. The pdf writer sometimes has problems printing in the background. This is do to problems caused by a combination of the modern memory manager, the print monitor, and the pdf printer driver itself. If a problem occur while printing in the background, just simply turn the modern memory manager off. The switch for the modern memory manager can be found in the memory control panel. That should clean up any difficulties.

Conclusion

Once the basics of working with html and pdf files are learned it is fairly easy to keep a web sight up to date. I suggest that a book like *Teach Yourself WebPublishing with HTML 3.0* should be picked up not only to learn how to write html code but also as a reference for even web veterans.

Appendix A - Directory Structure

www

www/booster

booster/images

booster/BoosterNetwork

booster/InterfaceChassis

booster/docs

linac/gifs

www/linac

linac/imageconf

linac/software

linac/hardware

software/macintosh

software/sun

software/locsys

hardware/samhold

hardware/srm

hardware/vmesys

hardware/DZero

hardware/genescript

hardware/ipcards

hardware/quickdig

irm/analogio

www/irm

irm/images

irm/binio

irm/misc_con

irm/database

irm/network

irm/overview

irm/refmat

irm/timer

irm/pdfFiles

Key

1) = main World Wide Web directory in under /home/florian/

2) = main sub directories.

3) = minor directories

4) = minor directory that contains image configuration files the linac director contains the image configurations for booster and irm main sub directories as well.

5) = main directories for images (mainly gifs)

6) = is or cantains a pdf directory which stores documents in pdf format